Plant leaf disease detection using deep learning techniques

Omkar Singh[1], Amit Kumar Pandey[2] , Anjali Rasal[3] , Abhishek Pandy[4]

[1]HOD(Department of Data Science) ,[2,3,4]PG Students

[1,3,4] Department of Data Science, [2]Department of IT,

Thakur College of Science and Commerce

Thakur Village, Kandivali(East), Mumbai-4000101, Maharashtra, India

Abstract :

Deep learning is a branch of artificial intelligence. In recent years, with the advantages of automatic learning and feature extraction, it has been widely used by academic and industrial circles. It has been widely used in image and video processing, voice processing, and natural language processing. At the same time, it has also become a research hotspot in the field of agricultural plant protection, such as plant disease recognition and pest range assessment, etc. The application of deep learning in plant disease recognition can avoid the disadvantages caused by the artificial selection of disease spot features, make plant disease feature extraction more objective, and improve the research efficiency and technology transformation speed. This review provides the research progress of deep learning technology in the field of crop leaf disease identification in recent years. In this paper, we present the current trends and challenges for the detection of plant leaf disease using deep learning and advanced imaging techniques. We hope that this work will be a valuable resource for researchers who study the detection of plant diseases and insect pests. At the same time, we also discussed some of the current challenges and problems that need to be resolved

1. Introduction :

Indian Economy is highly dependent on the agricultural productivity of the country. Grape is a very commercial fruit of India. It can easily be grown in all tropical, sub-tropical, and temperate climatic regions. India has different types of climate and soil in different parts of the country. This makes grapevines a major vegetative propagated crop with high socio-economic importance. The grape plant will cause poor yield and growth when affected by diseases. The diseases are due to viral, bacterial, and fungi infections which are by insects, rust, nematodes, etc.,

These diseases are judged by the farmers through their experience or with the help of experts through naked eye observation which is not an accurate and time-consuming process. Early detection of disease is then very much needed in agriculture and horticulture fields to increase the yield of crops. We have proposed a system that can detect and identify disease in the leaves of grape plants.

Keywords: Disease, India, types, grape, economy, agricultural productivity,

Leaf, CNN, KNN

2. Literature review :

The first research reference used was research from Yonatan Adi Winata et al., which discussed the research *using the Faster R-CNN method* with Inception-v2 architecture. Single Shot Detector (SSD) is also used in this research as a compliment. *The dataset they used in the research reference was the QUT FISH Dataset.* This research aimed to know the performance of the Faster R-CNN method against other object detection methods like SSD in fish species detection. *The accuracy of the Faster R-CNN reached 80.4%, far above the accuracy of the Single Shot Detector (SSD) Model with an accuracy of 49.2%.* The result of this research was a comparison of performance in fish species recognition between Faster R-CNN and SSD object detection methods [1].

Zhang, Huang, Huang, and Zhang discuss the crucial role played by plants in various domains such as agriculture, industry, medicine, and ecology. They emphasize the increasing threats to plant species brought about by global warming, biodiversity loss, urban development, and environmental degradation. Recognizing the importance of protecting plant species, the authors emphasize the need for efficient classification methods to identify and understand plants, particularly considering the vast number of both known and unknown species on Earth. The authors of the paper focus on plant species recognition through leaf analysis. They provide a comprehensive overview of existing methods, covering various aspects including plant leaf characteristics, public databases, and diverse recognition methods such as feature extraction, subspace learning, sparse representation, and deep learning. The paper emphasizes the simplicity and convenience of using leaves for recognition and aims to raise awareness about the importance of plant species identification. It serves as a valuable resource for beginners in the field, offering guidance and insights to foster a deeper understanding of plant species and contribute to their preservation.[2].

Jayashree Deka, Shakuntala Laskar, and Vikramaditya Baklial conducted a study on automated freshwater fish species classification using CNN. The major advantages of the deep CNN architecture over the manual supervised method are its self-learning and self-organized characteristics. The first layer in the CNN architecture is the image input primary layer that takes up 2-D and 3-D images as input. Deep-learning architectures, AlexNet and Resnet-50, to classify 20 indigenous freshwater fish species from the Northeastern parts of India. It incorporates 60 million parameters and 650,000 neurons. The architecture consists of 5 convolution layers, max-pooling layers, and 3 consecutive fully connected layers. The two models are fine-tuned for training and validation of the collected fish data. Performed networks are evaluated based on overall accuracy, precision, and recall rate. They reported the best overall classification accuracy, precision, and recall rate of 100% at a learning rate of 0.001 by the Resnet-50 model on their dataset and benchmark Fish-Pak dataset.[3].The study by Seppo Fagerlund discussed the use of support vector machines (SVM) for bird species recognition. This evaluated the performance of SVM on two different datasets. In the first dataset, where recognition testing was performed independently for each bird, the mixture model showed the best results among various generative representations, while the reference method employing MFCC parameters and nearest-neighbour classification also performed well. In the second dataset, where syllables were manually segmented, the SVM classifier demonstrated comparable performance to the reference method. Once again, the mixture model proved to be the most effective. The study concluded that SVM methods achieved equal or superior performance compared to reference methods. However, the study advised caution in directly comparing dataset results due to differences in species diversity and sound spectra. While the proposed method represented all syllables uniformly, the decision tree topology employed lacked consideration for sound relationships between species. Future work was suggested to explore feature weighting, which could improve accuracy, as seen in the Pygmy Owl case. Overall, the study highlighted SVM's potential for bird species recognition and suggested avenues for further refinement.[4]

Chen G, Han TX, He Z, Kays R, and Forrester T performed a study on species recognition for monitoring wild animals. They focused on deep convolutional neural networks (DCNN) and introduced a camera-trap dataset that comprises 20 North American species. This publicly available dataset includes 14,346 training images and 9,530 testing images, with color, Gray, and infrared images of resolutions ranging from 320 by 240 to 1024 by 768. To recognize different species, a bag-of-words model was used with image classification. This model divided images into blocks of 8 by 8, and treated them as "words". A histogram of occurrence counts for classification was used in this process. The code sizes (K) used in the BOW model were 1000, 2000, and 3000, which produced accuracies of 33.192%, 33.507%, and 33.485% respectively. It contrasted the BOW model and the DCNN algorithm on a dataset. The DCNN algorithm showed better species recognition accuracy of 38.315% compared to 33.507% for the BOW model. The DCNN algorithm could improve with more training data, despite the dataset's challenging nature. The DCNN algorithm could also identify ambiguous data for annotation, which would help reduce the burden on experts. Overall, the study demonstrated the effectiveness of the DCNN algorithm for species recognition in wildlife monitoring.[5]

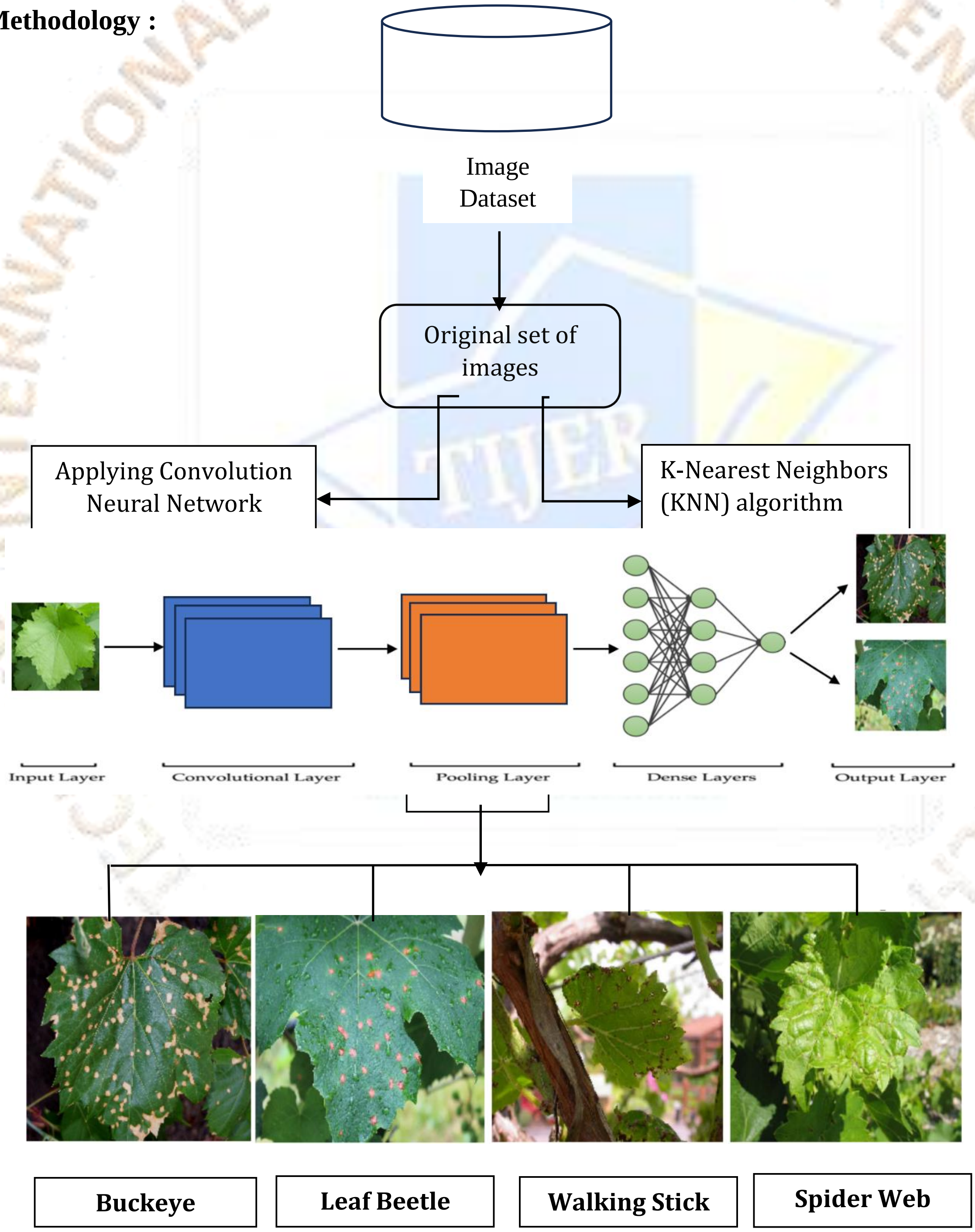
To enhance agricultural production and quality, this study focuses on automatic plant disease detection. To diagnose disorders associated with rice, it makes use of image processing, deep learning, machine learning, and meta-heuristic optimization. The study examines earlier research and offers insights into several methods for identifying diseases in rice plants. To recognize illnesses such as rice blasts, brown spots, leaf smut, tungsten, and sheath, a hybrid deep learning model is suggested.[6]

3. CNN :

A Convolutional Neural Network (CNN) has three layers: a convolutional layer, a pooling layer, and a fully connected layer. The figure shows all the layers together. CNN are fundamental component in the detection of leaf disease. Convolutional layers are the essence of CNNs. These layers apply a series of filters to the input image, scanning for specific features.

In the leaf disease detection project, these filters may produce an activation map by scanning the pictures several pixels at a time using a filter. Then reduce the amount of data created by the convolutional layer so that it is stored efficiently.

4. Methodology :



5. Result :

The detailed result of leaf disease detection is shown in the below table. Which shows the performance of the Supervised Model Based on Convolutional Neural Network (CNN) & K-Nearest Neighbors (KNN) throughout 25 epochs.

Epoch	Loss	Training Accuracy	Validation Accuracy
1/25	0.8490	0.4917	0.5250
2/25	0.7330	0.5250	0.5250
3/25	0.6836	0.5250	0.5250
4/25	0.6830	0.5250	0.5250
5/25	0.6996	0.4083	0.6500
6/25	0.6734	0.5500	0.5250
7/25	0.6651	0.5250	0.5417
8/25	0.6510	0.5583	0.7083
9/25	0.6349	0.6583	0.7250
10/25	0.6231	0.6667	0.5917
11/25	0.6097	0.6167	0.7083
12/25	0.5793	0.7167	0.7083
13/25	0.5482	0.7667	0.7667
14/25	0.5244	0.7417	0.7250
15/25	0.5401	0.6667	0.6917
16/25	0.5238	0.7333	0.7667
17/25	0.4814	0.7333	0.7917
18/25	0.4507	0.7833	0.8167
19/25	0.5107	0.7583	0.8083
20/25	0.4865	0.7417	0.8583
21/25	0.4668	0.7750	0.8500
22/25	0.4118	0.8667	0.8667
23/25	0.3739	0.8833	0.8250
24/25	0.3758	0.8333	0.8667
25/25	0.3387	0.8500	0.8333

With 89.17% accuracy rate was achieved using early stopping while Training the model on 25 epochs. The below figure show depicts the visualization of training and validation accuracy.

6. Conclusion :

Both the Convolutional Neural Network (CNN) and K-Nearest Neighbors (KNN) models were trained and evaluated over 25 epochs. The CNN model showed varying performance in terms of loss, training accuracy, and validation accuracy, while the KNN model did not show significant improvement in accuracy over the epochs.

The CNN model generally outperformed the KNN model, as indicated by the higher training and validation accuracies achieved by the CNN model across most epochs. The use of early stopping in training the model on 50 epochs suggests that the model's performance was monitored, and training was stopped when the model's performance on the validation dataset no longer improved.

7. Reference:

1. Fish Species Recognition with Faster R-CNN Adiwinata Y, Sasaoka A, Bayupati IA, Sudana O. Fish species recognition with Faster R-CNN Inception-v2 using QUT FISH dataset. Lontar Komputer: Jurnal Ilmiah Teknologi Informasi. 2020;11(3):144-54
2. Zhang S, Huang W, Huang YA, Zhang C. Plant species recognition methods using leaf image: Overview. Neurocomputing. 2020 Sep 30;408:246-72.
3. Automated Freshwater Fish Species Deka J, Laskar S, Baklial B. Automated Freshwater Fish Species Classification using Deep CNN. Journal of The Institution of Engineers (India): Series B. 2023 Apr 21:1-9.
4. Fagerlund S. Bird species recognition using support vector machines. EURASIP Journal on Advances in Signal Processing. 2007 Dec;2007:1-8.
5. Chen G, Han TX, He Z, Kays R, Forrester T. Deep convolutional neural network based species recognition for wild animal monitoring. In2014 IEEE International Conference on image processing (ICIP) 2014 Oct 27 (pp. 858-862). IEEE.
6. Tholkapiyan M, Aruna Devi B, Bhatt D, Saravana Kumar E, Kirubakaran S, Kumar R. Performance Analysis of Rice Plant Diseases Identification and Classification Methodology. Wireless Personal Communications. 2023 May;130(2):1317-41.